PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 17/27, 17/28

A1

(11) International Publication Number: WO 99/50759

(43) International Publication Date: 7 October 1999 (07.10.99)

(21) International Application Number: PCT/US99/06935

(22) International Filing Date: 29 March 1999 (29.03.99)

(30) Priority Data:

60/080,030	30 March 1998 (30.03.98)	US
09/281,996	29 March 1999 (29.03.99)	US

(71) Applicant: WORLDFREE.NET, INC. [US/US]; Suite 407, 9[illegible] Wilshire Boulevard, Beverly Hills, CA 90211 (US).

(72) Inventor: KIRCHMAN, Kevin, A., P.; 291 S. La Cienega Boulevard #107, Beverly Hills, CA 90211 (US).

(74) Agents: HOLLINGER, Joseph, K. et al.; Graham & James LLP, 600 Hansen Way, Palo Alto, CA 94304-1043 (US).

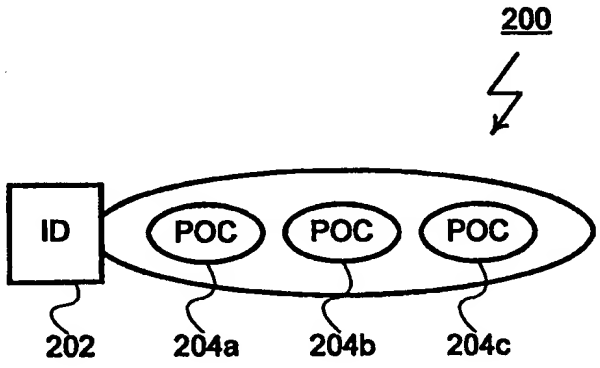
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD AND APPARATUS FOR USING IDEAS AND CONCEPTS WITHIN COMPUTER PROGRAMS

(57) Abstract

A method and apparatus that allows computer programs to define ideas and concepts symbolically is provided. The method and apparatus include a grammar that may be used to represent any concept. Sentences are parsed, using this grammar, into their component parts. As part of the parsing process, each word is compared to the contents of a dictionary database. The dictionary database and a set of tense-mood tables are used to identify individual words as concepts, entities, actions or qualifiers. The parsing process creates a data structure (200) for each sentence. The data structure organizes the sentence into its component parts, such as an ID field (202) and POC fields (204). The data structures for different sentences can be compared to determine matching or similarity. The data structures can also be processed to accomplish more advanced ends, such as reasoning systems or expert systems.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Method And Apparatus For Using Ideas and Concepts Within Computer Programs

FIELD OF THE INVENTION

This application relates generally to the use of human concepts and ideas within interactive computer applications. More specifically, the present invention includes a method and apparatus that allows ideas to be defined symbolically for processing and matching with other ideas.

BACKGROUND OF THE INVENTION

The inability of computer applications to process human ideas and concepts has many disadvantages. As an example, consider the case where a person wants an answer for the question: "what is the largest dinosaur?" In today's world, it is increasingly common to look for the answer to such questions on the World Wide Web. This type of searching is typically done by formulating and submitting queries to the many search engines available on the web. Queries of this type are formulated using keywords along with connectors and qualifiers.

Unfortunately, the use of the queries means that the search engine has no real way of knowing what the user is actually searching for. The search engine never sees or appreciates the question: "what is the largest dinosaur?" As a result, the search engine is limited in the type of search it can perform. It can look for web sites that contain the specified keywords modifies by the specified connectors and qualifiers. It cannot compare the content of web sites to determine if there is a match for the largest dinosaur question (the search engine is also unaware of the meaning of that content). This makes the searching process rather hit and miss. In many

cases, search engines users scan many sites to find the content they are looking for. In other cases, the desired content is never found.

This failure to understand ideas and concepts also complicates the interaction between users and search engines (and other applications). The user is forced to translate the question "what is the largest dinosaur?" into an acceptable query. This is much more difficult than simply typing in the original question. The arcane nature of these queries often means that the user will make several attempts before formulating an acceptable query.

The web searching example is, of course, a highly limited example. The failure to understand ideas and concepts extends to an almost unlimited number of different applications in an almost unlimited number of ways. For these reasons, a need exists for solutions that allow computers and their applications to more fully understand and process human ideas and concepts.

SUMMARY OF THE INVENTION

An embodiment of the present invention includes a method and apparatus that allows computer programs to define ideas and concepts symbolically. The symbolic ideas and concepts may be processed and matched with other symbolically defined ideas.

The present invention defines a grammar that may be used to represent any concept. Proper concepts (the equivalent of complete sentences) form the highest level of this grammar. Entities, and actions (the equivalent of individual words) form the lowest level of the grammar. Proper concepts are defined as recursive combinations of concepts, entities, actions and qualifiers.

Sentences are parsed, using this grammar, into their component parts. As part of the parsing process, each word is compared to the contents of a dictionary database. The dictionary database and a set of tense-mood tables are used to identify individual words as concepts, entities, actions or qualifiers.

The parsing process creates a data structure for each sentence. The data structure organizes the sentence into its component parts. The data structures for different sentences can be compared to determine matching or similarity. The data structures can also be processed to accomplish more advanced ends, such as reasoning systems or expert systems.

Advantages of the invention will be set forth, in part, in the description that follows and, in part, will be understood by those skilled in the art from the description herein. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims and equivalents.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments of the invention and, together with the description, serve to explain the principles of the invention.

Figure 1 is a block diagram of a host computer system shown as an exemplary environment for an embodiment of the present invention.

Figure 2 is a block diagram of a concept data structure as used by an embodiment of the present invention.

Figure 3 is a block diagram of a concept represented as a combination of two concept data structures.

Figure 4 is a block diagram of a possible implementation of the concept data structure as used by an embodiment of the present invention.

Figure 5 is a second block diagram of a concept represented as a combination of two concept data structures.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Reference will now by made in detail to preferred embodiments of the invention, examples of which are illustrated in the accompanying drawings. Wherever convenient, the same reference numbers will be used throughout the drawings to refer to the same of like parts.

ENVIRONMENT

In Figure 1, a computer system 100 is shown as a representative environment for the present invention. Structurally, computer system 100 includes a processor, or processors 102, and a memory 104. An input device 106 and an output device 108 are connected to processor 102 and memory 104. Input device 106 and output device 108 represent a wide range of varying I/O devices such as disk drives, keyboards, modems, network adapters, printers and displays. Each node 102 may also includes a disk drive 110 of any suitable disk drive type (equivalently, disk drive 110 may be any non-volatile mass storage system such as "flash" memory).

GRAMMAR

An embodiment of the present invention includes a method and apparatus that allows computer programs to define and manipulate concepts symbolically. The symbolic concepts may be processed and matched with other symbolically defined concepts. Part of the method and apparatus of the present invention are a grammar that defines how concepts and ideas can be symbolically represented. A basic premise of this grammar is the notion that any concept may be representing in cannonical form. One possible cannonical form, referred to as the general form of a concept, is:

$C \equiv E(Q)\ A(Q)\ E(Q)$

where

$C \equiv$ *Concept (Complex)*

$E \equiv$ *Concept of Entity*

$E \equiv$ *Concept of Action*

$Q \equiv$ *Concept of Qualification*

The symbol '≡' establishes identity.

Concept entities and concept actions are words, concept qualifiers are further defined as any subset of:

$Q \equiv E(Q)\ A(Q)\ E(Q)$ or

$Q \equiv Q\ Q\ Q\ \ldots$

where the concept of action can be a noun in the statement. An indirect object is considered as qualifying the concept of action. It is important to recognize that subordinate, coordinate and independent clauses are being represented as qualifying the entity or action that they modify and can be considered as limitations upon the sets of referents of those terms.

Subordinating qualifiers are of the form:

Q(Q)

while coordinating qualifiers are of the form:

QQ

A Proper Concept, the equivalent of a complete sentence, is either one of the forms:

E(*Q*) *A*(*Q*)

E(*Q*) *A*(*Q*) *E*(*Q*)

∞*C* ∞*C*

where Q is optional and

∞ ≡ *Concept of Connectivity*

The last concept above is a compound concept, representing the combination of two or more complex concepts with optional connectives, such as 'if', 'and' or 'that'. Compound concepts include conditional statements, intentional statements and lists.

EXAMPLES OF CONCEPTS REPRESENTED SYMBOLICALLY

The following examples serve to demonstrate the use of the previously defined grammar. To use the grammar, words within a sentence being represented are identified with reference to a dictionary database (within the following examples, this is represented by subscripted letters). The ending of each word is compared to a tense-mood table and categorized as one of E, A or Q.

Example 1: The wood was burnt.

E: wood the

A: burn past

E:

$$E_p (Q_{in}) A_x (Q_{rt})$$

The general form of this concept is: $C \equiv E_p (Q_{in}) A_x (Q_{rt})$

In this case the subject 'wood' is not qualified with respect to quantity because no quantity is stated.

Example 2: Those men skillfully produced blue houses from that red wood.

E: human male plural

A: produced imperfect skillfully from wood that

E: hose plural blue

$$C \equiv E_h\,[Q_{in}\,Q_{ge}\,Q_{in}]\;A_x\,[Q_{rt}\,Q_{rq}\,Q_{so}\,[E_p\,(Q_{in}\,Q_{qu})]]\;E_p\,[Q_{nu}\,Q_{qu}]$$

Example 3: It is raining.

The indefinite pronoun here must be made explicit. Rephrased: The weather is the state of raining.

$$C \equiv E_p\,(Q_{in})\;A_x\,(Q_{qu})\;E_p\,(Q_{st})$$

Example 4: The pretty ladies danced gracefully before the guests.

$$C \equiv E_h\,[Q_{ge}\,Q_{in}\,Q_{nu}\,Q_{rq}]\;A_h\,[Q_{rt}\,Q_{rq}\,Q_{rp}\,[E_h\,Q_{ca}\,Q_{in}\,Q_{nu}]]$$

Example 5: Fred, the builder who lives in the woods, fabricated a cart for hauling his tools.

$$C \equiv E_h\,[Q_{in}\,[Q_{in}\,Q_{ca}\,[A_h\,[Q_{rt}\,Q_{rp}\,[Ep\ Qin]]]]]$$
$$A_h\,[Q_{rt}\,Q_{mo}\,[A_h\,[Q_{rt}\,E_p\,[Q_{nu}\,Q_{po}]]]]$$
$$E_p$$

Example 6: He cooled the hot metal by pouring cold water on it.

$$C \equiv E_h\,[Q_{in}\,Q_{co}]$$
$$A_p\,[Q_{rt}\,Q_{ma}\,[A_h\,[Q_{rt}\,Q_{rp}\,[E_p\,[Q_{in}\,Q_{rq}]]]$$
$$E_p\,[Q_{rq}]]]$$
$$E_p\,[Q_{in}\,Q_{rq}]$$

Example 7: The present king of France is bald.

$$C \equiv E_h\,[Q_{ca}\,[Q_{so}]\,Q_{in}\,Q_{rt}]\,A_x\,E_h\,Q_{qu}]$$

Example 8: 'It was to their native law that sixteenth century Scottish Jurists applied most successfully the humanistic methods first developed on the continent in the study of Roman Law'

(sixteenth century Scottish Jurists)

(applied most successfully) (to their native law)

(the humanistic methods which were first developed on the continent in the study of Roman Law)

$$C \equiv E_h\,[Q_{ca}\,Q_{so}\,Q_{rt}\,Q_{no}]$$
$$Ax\,[Q_{ro}\,[Q_{rt}]\,Q_{re}\,[E_m\,[Q_{so}\,Q_{po}]]]$$
$$Em\,[Q_{in}\,Q_{nu}\,Q_{qu}\,A_x\,[Q_{rt}\,Q_{rp}\,[E_p\,[Q_{in}]]\,Q_{ma}\,[A_h\,[E_m\,Q_{so}]]]]$$

The word 'law' refers to as concepts (though recorded as physical entities) and thus are Mental Entities (E_m). The word 'methods' also refers to concepts. The words 'were first' are considered together as a claim of Relative Temporality,

although 'first' might also be considered as a qualification of Source. The words 'in the study of' are considered to 'by means of studying'.

Concepts and ideas are recursively parsed using the grammar. During the parsing process, a data structure is created to represent the concept or idea being parsed. The data structure is a tree-shaped hierarchy. Each node in the tree is a data structure of the form shown as concept data structure 200 of Figure 2. Concept data structure 200 includes an ID 202 and three POC fields 204a through 204c. ID 202 allows each concept data structure 200 to be uniquely identified. The inclusion of three POC fields 204 follows the general form for concepts described above. Each POC field 204 can represent an action or entity. POC fields can also point to other concept data structures 200. This allows POC fields 204 to point to other concepts. This is shown, for example, in Figure 3 where the concept "That the stock went up is a good sign" is shown as a combination of two concept data structures 200a and 200b.

Figure 4 shows one possible implementation for concept data structure 200. Concept data structure 200 of Figure 4 includes fields for ID, Word, Wordtype, Q1, Q2, Q3, Q4, Qualified List Pointer, Subordinate Concept Pointer, PrevPointer, PostPointer, Program and Referents. The implementation of Figure 4 is intended to be representative. Other implementations may be used without departing from the spirit of the present invention.

Figure 5 shows a concept that has been parsed into the tree data structure described in regard to Figures 2 though 4. In this case, the concept includes an initial entity ("e1word") followed by an action ("a1word) followed by a second entity

("e2word). The initial entity is qualified by a second concept ("Cword"). The second concept is composed of still more concepts.

Data structures of the type described in Figures 2 through 5 are recursively constructed for each proper concept (sentence) that is represented using the method of the present invention. Once represented as a data structure of this type, a sentence or concept can be manipulated within computer programs. In particular, matching between concepts can be determined by comparing their data structures. For matching, data structures are examined to determine if they have the same (or similar structures). The contents of the data structures are also compared to determine sameness or similarity.

Other embodiments be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope of the invention being indicated by the following claims and equivalents.

WHAT IS CLAIMED IS:

1. A method for creating a symbolic representation of a sentence, the method comprising:

recursively parsing the sentence to isolate one or more concepts, each concept being a cannonical arrangement of entities, actions and qualifiers;

creating concept data structures to represent each respective concept that is isolated; and

linking the concept data structures to form a hierarchical data structure representing the sentence.

2. A method as recited in claim 1 wherein each concept is a cannonical arrangement of the form E(Q) A(Q) E(Q) where E is an entity, A is an action and Q is a qualifier.

3. A method as recited in claim 2 wherein each qualifier is a list of qualifiers or is a cannonical arrangement of the form E(Q) A(Q) E(Q).

4. A method as recited in claim 1 further comprising the step of comparing at least one word in the sentence to the contents of a dictionary database to determine if the word is a qualifier, entity or action.

5. A method as recited in claim 4 wherein the dictionary database includes tense-mood tables.

6. A method for creating a symbolic representation of a sentence, the method comprising:

using a dictionary database to categorize words in the sentence as qualifiers, entities or actions;

grouping the qualifiers entities and actions into concepts, each concept being a cannonical arrangement of entities, actions and qualifiers;

creating concept data structures to represent each respective concept that is created; and

linking the concept data structures to form a hierarchical data structure representing the sentence.

7. A method as recited in claim 6 wherein each concept is a cannonical arrangement of the form E(Q) A(Q) E(Q) where E is an entity, A is an action and Q is a qualifier.

8. A method as recited in claim 7 wherein each qualifier is a list of qualifiers or is a cannonical arrangement of the form E(Q) A(Q) E(Q).

9. A method as recited in claim 6 further comprising the step of comparing at least one word in the sentence to the contents of a dictionary database to determine if the word is a qualifier, entity or action.

10. A method as recited in claim 9 wherein the dictionary database includes tense-mood tables.

11. A computer program product comprising a computer usable medium having computer readable code embodied therein, the computer readable program

code devices configured to cause a computer system to perform a method for creating a symbolic representation of a sentence, the method comprising:

rendering the image portion at a higher resolution to create a texture;

retrieving the rendered image portion as a texture;

filtering the texture to produce a minified texture; and

applying the minified texture to a quadrilateral within the graphics image.

12. A computer program product as recited in claim 11 wherein each concept is a cannonical arrangement of the form E(Q) A(Q) E(Q) where E is an entity, A is an action and Q is a qualifier.

13. A computer program product as recited in claim 12 wherein each qualifier is a list of qualifiers or is a cannonical arrangement of the form E(Q) A(Q) E(Q).

14. A computer program product as recited in claim 11 wherein the method further comprises the step of comparing at least one word in the sentence to the contents of a dictionary database to determine if the word is a qualifier, entity or action.

15. A computer program product as recited in claim 14 wherein the dictionary database includes tense-mood tables.

Fig. 1

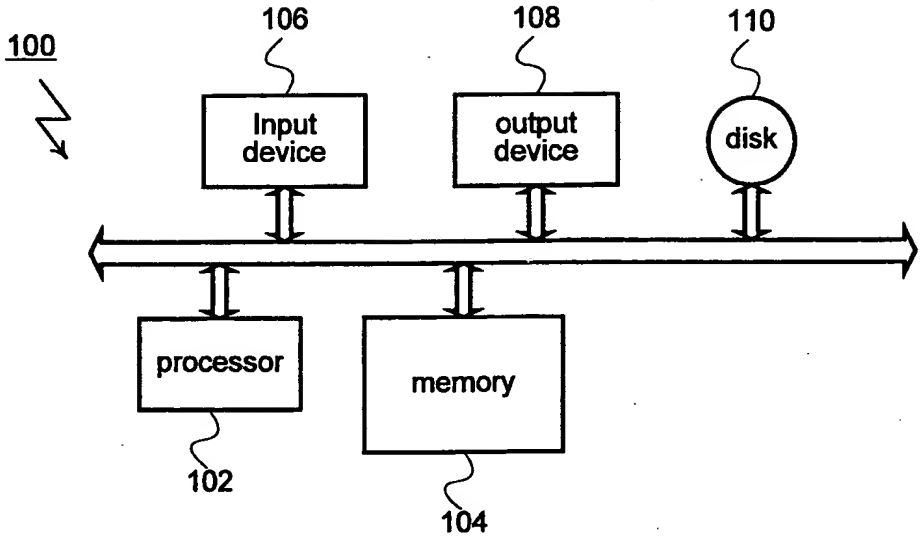


Fig. 2

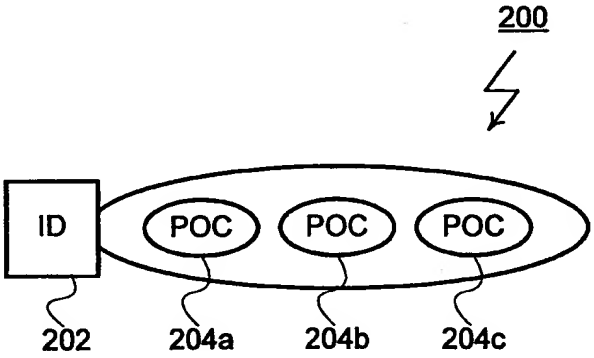
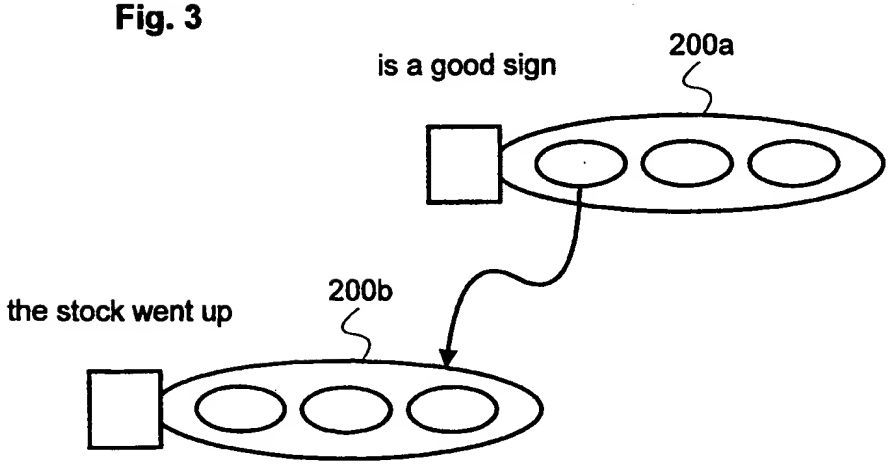


Fig. 3



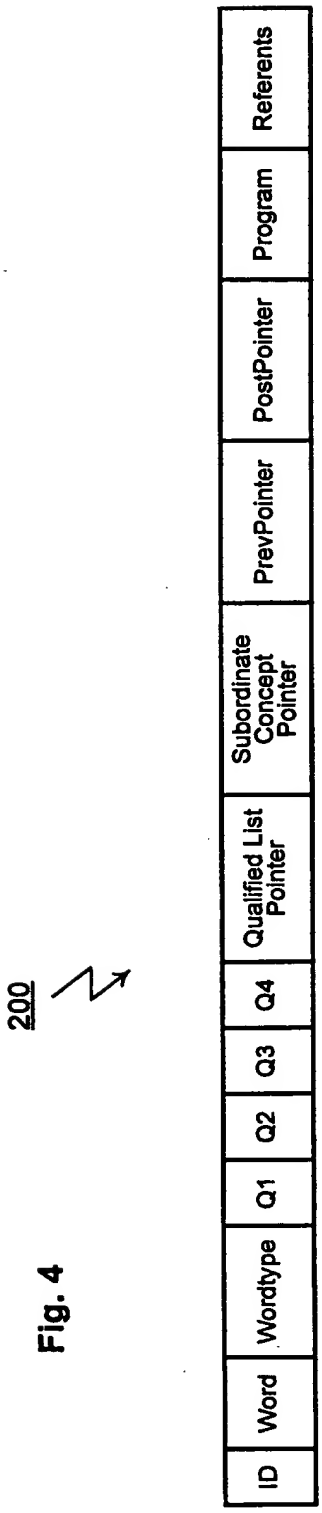

200
ID
Word
Wordtype
Q1
Q2
Q3
Q4
Qualified List Pointer
Subordinate Concept Pointer
PrevPointer
PostPointer
Program
Referents

Fig. 4

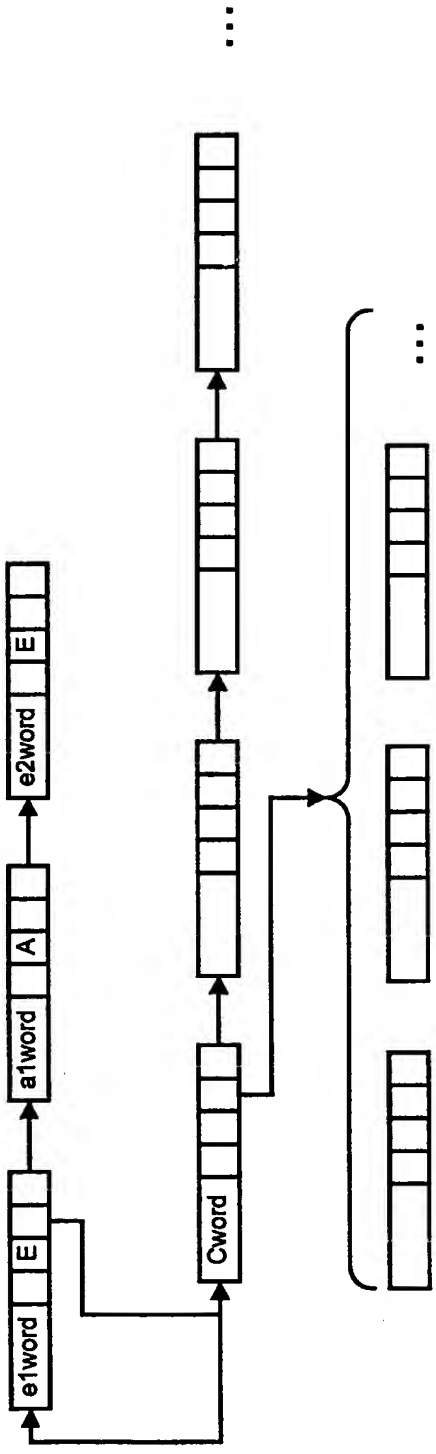

e1word
E
a1word
A
e2word
E
Cword

Fig. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/06935

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :G06F 17/27, 17/28
US CL :704/9
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/9, 1, 10; 707/2, 3, 5, 6, 100, 102, 104, 530, 531, 532; 434/118, 156, 167, 169; 706/55

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,197,005 A (SCWARTZ et al.) 23 March 1993, abstract; figs. 1-5 & 7; col. 4, line 9 to col. 5, line 23; and col. 5, line 51 to col. 22, line 32	1-15
A	US 5,377,103 A (LAMBERTI et al.) 27 December 1994, abstract; figs. 1-3; and col. 3, line 37 to col. 8, line 19	1-15
Y	US 5,454,106 A (BURNS, et al.) 26 September 1995, abstract; figs. 1-7; col. 1, line 15 to col. 3, line 30; and col. 3, line 59 to col. 11, line 1	1-15
A, P	US 5,794,050 (DAHLGREN et al.) 11 August 1998, abstract; figs.1-5 & 7; 5, line 7 to col. 6, line 12; and col. 6, line 36 to col. 16, line 4	1-15

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 19 MAY 1999	Date of mailing of the international search report 15 JUN 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer James R. Matthews JOSEPH THOMAS Telephone No. (703) 308-9588

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/06935

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	US 5,878,385 A (BRALICH et al.) 02 March 1999, abstract; figs. 1-3, 7, & 11; col. 1, line 5 to col. 5, line 36; col. 6, line 13 to col. 14, line 37; col. 46, line 65 to col. 48, line 45; and col. 51, line 59 to col. 57, line 21	1-15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/06935

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS

search terms: parse/parser/parsing, canonical, heirarchy/heirarical, grammar/grammatical,data structure